NOTES FOR 27 SEPT (WEDNESDAY)

1. How to come up with submanifolds ?

How do you come up with submanifolds of $\mathbb{R}^2$? One example is the circle $x^2 + y^2 = 1$. It is a "level set" of a function, i.e., $S^1 = f^{-1}(1)$ where $f(x, y) = x^2 + y^2$. Likewise, more generally, if you have a smooth function $f : M \to N$, then one way to hope to come up with submanifolds of M is to take level sets of f, i.e., $f^{-1}(q)$ where $q \in N$. Unfortunately, they may not always be submanifolds. Even if they are submanifolds, they may not have the "correct" dimension. (After all, suppose N has dimension n. Then if you look at all $p \in M$ such that $f(p) = q$, then you are imposing n conditions. There are m free parameters. So you should $m - n$ free parameters left. So you should hope that $f^{-1}(q)$ is an $m - n$ dimensional manifold.) Examples and counterexamples :

(1) If $f(x, y) = x^2 + y^2$, then $f^{-1}(0) = (0, 0)$. This is a submanifold but it has dimension 0 (instead of 1).
(2) Likewise, for the same f, $f^{-1}(-1) = \phi$. An empty set is a submanifold of whatever dimension you want.
(3) Suppose $g(x, y) = x^2 - y^2$. Now $g^{-1}(0)$ consists of two straight lines $y = \pm x$. The level set is a submanifold of dimension 1 everywhere except at the origin.
(4) Suppose $h(x, y, z) = (x^2 + y^2 + z^2, x + y + z)$. Then $h^{-1}(1, \sqrt{3})$ is a point $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, i.e., a submanifold of dimension 0 (instead of 1).

What is it that makes some preimages (or level sets) submanifolds and some not ? The crucial point here is the implicit function theorem. If you have a function $f : \mathbb{R}^m \to \mathbb{R}^n$ such that $Df(p)$ is a *surjective* linear map, then denoting $f(p) = q$, the set of points $x \in \mathbb{R}^m$ near p satisfying $f(x) = q$ (i.e. the preimage of q near p) is a submanifold of dimension $m - n$. Let us look at the problematic examples above.

(1) The derivative matrix of f is a row matrix, $Df = [2x, 2y]$. When $x = y = 0$, this derivative is 0, i.e., it is not a surjective linear map. So there is no surprise that $f^{-1}(0, 0)$ is not what we want.
(2) Obvious.
(3) $Dg = [2x, -2y]$ which is $[0, 0]$ at the origin. Not a surjective linear map from $\mathbb{R}^2 to \mathbb{R}$.
(4) For h, the derivative matrix is a 2×3 matrix.

$$Dh = \begin{bmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{bmatrix} \tag{1.1}$$

Now clearly when $x = y = z$, this matrix does not have rank 2 (i.e. full rank). So the derivative is not surjective as a matrix from $\mathbb{R}^3$ to $\mathbb{R}^2$ at that point. No surprise that problems occur there.

More generally, given a map $f : M \to N$, $f^{-1}(q) \subset M$ is an embedded submanifold of dimension $m - n$ if q is regular value, i.e., after choosing coordinates, the derivative map is a surjective linear map at all points in the pre-image of q. In the above examples, q was a critical value, i.e., there are *some* points in the pre-image where the derivative map is not surjective.

A small warning : if q is a regular value, then $f^{-1}(q)$ is a submanifold of dimension $m - n$. But the converse need not be true. Just because q is not a regular value does not mean that $f^{-1}(q)$ cannot be a submanifold of the correct dimension. Here is a stupid example : Take $f(x, y) = x^2$. Sard's theorem states that the set of critical values has measure 0. In other words, for almost all $q \in N$, either

(1) $f^{-1}(q)$ is empty, or
(2) $f^{-1}(q)$ is a submanifold of M of dimension $m - n$.

In our midterm, the question was (roughly) if Z is a submanifold of N of dimension z, then under a technical assumption ("f is transverse to Z") is $f^{-1}(Z)$ a submanifold of M ? How can we solve this problem ? The hint indicated was that locally there exists a coordinate system on N such that $y^1 = y^2 = \ldots = y^{n-z} = 0$ is Z, i.e., locally, you can define Z as the preimage/level set of $g : U \subset N \to \mathbb{R}^{n-z}$ given by $g(y^1, \ldots, y^n) = (y^1, \ldots, y^{n-z})$. Now $f^{-1}(Z) = (g \circ f)^{1-}(0)$. So at least locally, we have a map $f^{-1}(U) \subset M \to \mathbb{R}^{n-z}$ given by $h = g \circ f$ such that $f^{-1}(Z) \cap f^{-1}(U)$ is a level set of h. So this level set is locally a submanifold of dimension $m - (n - z)$ provided 0 is a regular value, i.e., for every point in the preimage, Dh is surjective (after choosing coordinates on M to find Dh). To do this, we need use the assumption of transversality, i.e., $f_*(T_pM) + T_{f(p)}Z = T_{f(p)}N$. For us to use this, we need to remember what a pushforward is. We will recall that in more detail in a moment. But at a very simple level, the pushforward should be thought of as the derivative matrix after choosing coordinates in the domain and the target. All transversality is saying is that the image of $n \times m$ matrix $Df(p) = \begin{bmatrix} \frac{\partial f^1}{\partial x^1} & \frac{\partial f^1}{\partial x^2} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f^{n-z}}{\partial x^1} & \frac{\partial f^{n-z}}{\partial x^2} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$ plus the vector space spanned by vectors of the form $(0, 0, 0 \ldots, 0, 1, 0 \ldots)$ (where the 1 occurs after $n - z$ slots) is all of $\mathbb{R}^n$ for all $p \in f^{-1}(Z)$. In other words, the image of Df must at least contain the first $n - z$ basis vectors of $\mathbb{R}^n$ (it may be bigger though). This means that at least $n - z$ columns of $Df(p)$ should be linearly independent for all $p \in f^{-1}(Z)$ and be independent of the last z basis vectors. Thus the first $n - z \times m$ part of Df which is nothing but Dg should have full rank.

2. Tangent and cotangent bundles

Firstly, a vector bundle $V \to M$ is a manifold with an onto map to M such that every preimage is a vector space, and locally, it is isomorphic to $U \times \mathbb{R}^r$. Recall that every vector bundle is isomorphic to a weird quotient using transition functions. Transition functions are smooth maps $g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(r, \mathbb{R})$ satisfying

(1) $g_{\alpha\alpha} = Id$
(2) $g_{\alpha\beta} = g_{\beta\alpha}^{-1}$
(3) $g_{\alpha\beta} g_{\beta\gamma} g_{\gamma\alpha} = Id$.

There are lots of examples of vector bundles. Two important *examples* of vector bundles associated to every manifold M are its tangent TM and cotangent T^*M bundles. The cotangent bundle T^*M was constructed as a weird quotient with transition functions $[g_{\alpha\beta}]_{ij} = \frac{\partial x^i_\beta}{\partial x^j_\alpha}$. The tangent bundle was constructed by putting a topology and a vector bundle structure on the set $\cup_p T_pM$. Interestingly enough, if you write the transition functions for the tangent bundle, they are $[h_{\alpha\beta}]_{ij} = \frac{\partial x^i_\alpha}{\partial x^j_\beta}$. Note that

$\sum_{j=1}^{n}[g_{\alpha\beta}]_{ij}[h_{\alpha\beta}]_{kj} = \delta_{ij}$. This means that $h_{\alpha\beta} = (g_{\alpha\beta}^{-1})^T$.

This reminds us of a familiar linear algebra concept : Suppose V is a vector space and $e_1, \ldots, e_n$ is a basis. Then every vector is of the form $v = \sum_{i=1}^{n} v_i e_i$. Given a basis of V, there is a "natural" basis for V^* given by linear functionals ω_i defined by $\omega_i(e_j) = \delta_{ij}$. So every linear functional w is of the form $w = \sum_i w_i \omega_i$. Note that the pairing $w(v)$ is $w(v) = \sum_i w_i v_i$. In other words, the expression $\sum_i w_i v_i$ is independent of the basis chosen. Moreover, suppose we choose a new basis $\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n$ for V. Then the new components for $v = \sum_i \tilde{v}_i \tilde{e}_i$ are related to the old by an invertible matrix, $\tilde{v}_i = \sum_j P_{ij} v_j$. What about the corresponding linear functional components $\tilde{w}_i$? Note that since $\sum_j w_j v_j = \sum_i \tilde{w}_i \tilde{v}_i$, we see that $\sum_j w_j v_j = \sum_i \sum_j P_{ij} v_j \tilde{w}_i$. Thus, $\tilde{w} = (P^{-1})^T w$.

In other words, every cotangent space is dual to the corresponding tangent space. Given a 1-form $\omega = \sum_{i=1}^{n} \omega_i dx^i$, and a vector field $X = \sum_{i=1}^{n} X_i \frac{\partial}{\partial x^i}$, we can form a scalar $\omega(X) = \sum_{i=1}^{n} \omega_i X_i$ (which is independent of the coordinate system chosen). The cotangent bundle is the "dual bundle" of the tangent bundle.

More generally, suppose we are given a vector bundle V with transition functions $g_{\alpha\beta}$, we can form a new set of matrix-valued functions $h_{\alpha\beta} = ([g_{\alpha\beta}]^{-1})^T$. You can easily verify that $h_{\alpha\beta}$ satisfy the requirements to be the transition functions of another vector bundle. That vector bundle is called the "dual" bundle to V and is denoted as V^*. Every fibre of V^* is dual to every fibre of V. Given a section s of V and t of V^*, we can get a function using $t(s)$.

Coming back to the tangent and cotangent bundles, given a basis $\frac{\partial}{\partial x^i}$ of the tangent bundle, the corresponding dual basis is none other than dx^i which satisfies $dx^i(\frac{\partial}{\partial x^j}) = \delta_{ij}$.

3. What the heck is a pushforward f_* ? What is df ? What is a pull-back f^*? What is the difference ?

Given a smooth function $f : \mathbb{R}^2 \to \mathbb{R}$, what is its derivative Df? commonly, in multivariable calculus, it is a row matrix $[\frac{\partial f}{\partial x} \frac{\partial f}{\partial y}]$. The directional derivative $D_v f = \frac{\partial f}{\partial x} v_1 + \frac{\partial f}{\partial y} v_2 = [Df] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. So the derivative Df can be thought of as a linear functional $\mathbb{R}^2 \to \mathbb{R}$ given by $Df(v) = [Df][v]$. Alternatively, we can think of the gradient ∇f as a column vector $\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$. But usually this is not done in any class on multivariable calculus.

More generally, if you have $f : \mathbb{R}^m \to \mathbb{R}^n$, the derivative matrix Df is an $n \times m$ matrix consisting of $\begin{bmatrix} Df_1 \\ Df_2 \\ Df_3 \\ \vdots \end{bmatrix}$. It takes a vector from $\mathbb{R}^m$ and spits out a vector in $\mathbb{R}^n$.

Given a smooth map $f : M \to N$, we would like to know what the "derivative" of this map is. Unfortunately, we can't simply take coordinates and write down the derivative matrix $Df : M \to \mathbb{R}^n$ because such an object will depend on the coordinates chosen. However, given a tangent vector $v \in T_pM$, we can produce another tangent vector $w \in T_{f(p)}N$ in a natural manner. Indeed, suppose v is the equivalence class of a curve $[c(t)]$, then $w = f_* v = [f \circ c(t)]$. Alternatively, if v is a derivation, then $w = f_* v$ is a derivation that acts on a function g as $w(g) = v(g \circ f)$. At the level of coordinates, suppose $v = \sum_i v_i \frac{\partial}{\partial x^i}$, then $f_* v = \sum_{i,j} v_i \frac{\partial f^j}{\partial x^i} \frac{\partial}{\partial y^j}$. So indeed, the derivative matrix in coordinates IS the pushforward in coordinates. No difference!